

# Computer Vision

CSE/ECE 576

Linda Shapiro

Professor of Computer Science & Engineering  
Professor of Electrical & Computer Engineering

# Course Information

- Time:
  - MW: 1:30-2:50
- Location:
  - ECE 037
- Contact:
  - shapiro@cs.uw.edu
- TAs:
  - Sitong Liu
  - [sitonl2@uw.edu](mailto:sitonl2@uw.edu)
  - Nishat Khan
  - [nkhan51@uw.edu](mailto:nkhan51@uw.edu)
  - Yuekai Xu (code problems only)
  - [yuekai@uw.edu](mailto:yuekai@uw.edu)
- Website:
  - <https://courses.cs.washington.edu/courses/cse576/26sp/>



# Topics

- Introduction
- Color and Texture
- Image Coordinates, Transforms, and Resizing
- Filters and Convolutions
- Edges and Lines
- Interest Operators, Image Matching, Image Stitching
- Face Detection/Recognition
- Machine Learning Overview including Neural Nets
- Object Detection and Recognition with ML
- Convolutional Neural Networks
- Self-Supervised Learning
- Motion/Optical Flow
- Stereo and 3D Depth Perception

# Grading (tentative)

- Six regular assignments (75%)
- One Course Project (25%)
- NO EXAMS (Yay!)

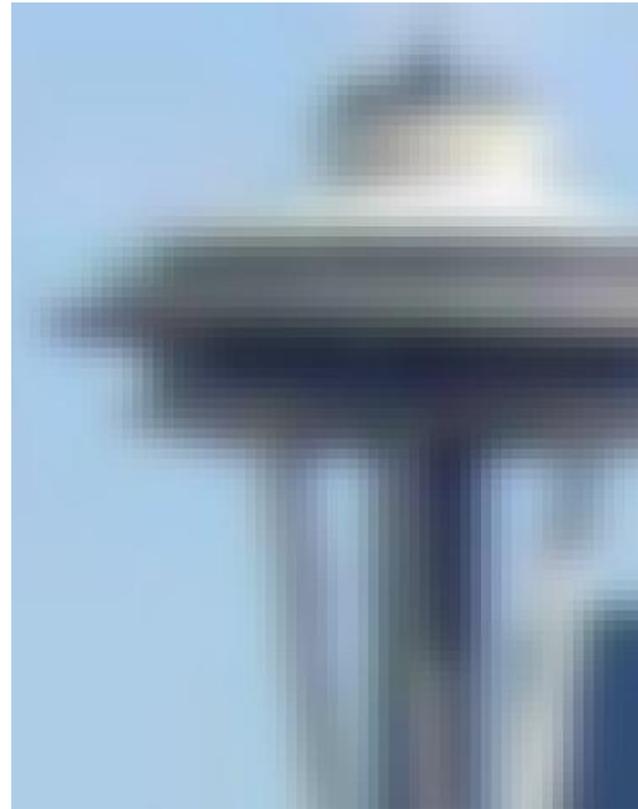
# Assignments

- Build a vision library from the ground up
- Recently converted to Python; one in PyTorch
- Play with advanced tools, neural networks
  
- Beginning: lots of skeleton code, explanations
- End: less guidance, more experimentation

# Assignment 1: Fun with Color



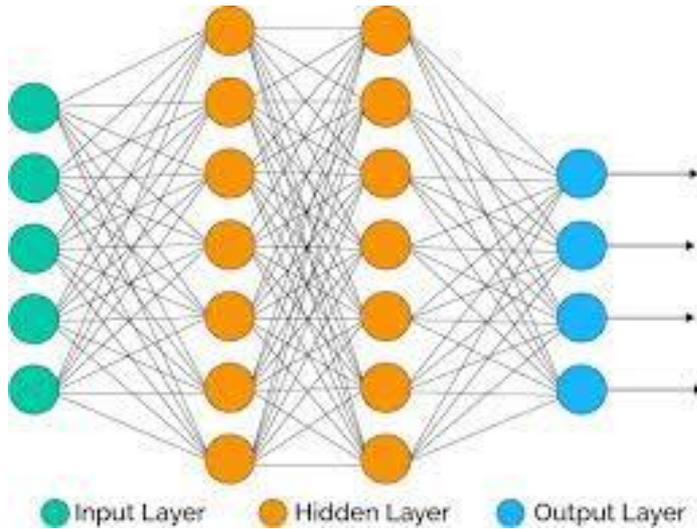
# Assignment 2: Image Resizing and Filtering



# Assignment 3: Panorama Stitching



# Assignment 4: Neural Networks



Here are the classes in the dataset, as well as 10 random images from each:

**airplane**



**automobile**



**bird**



**cat**



**deer**



**dog**



**frog**



**horse**



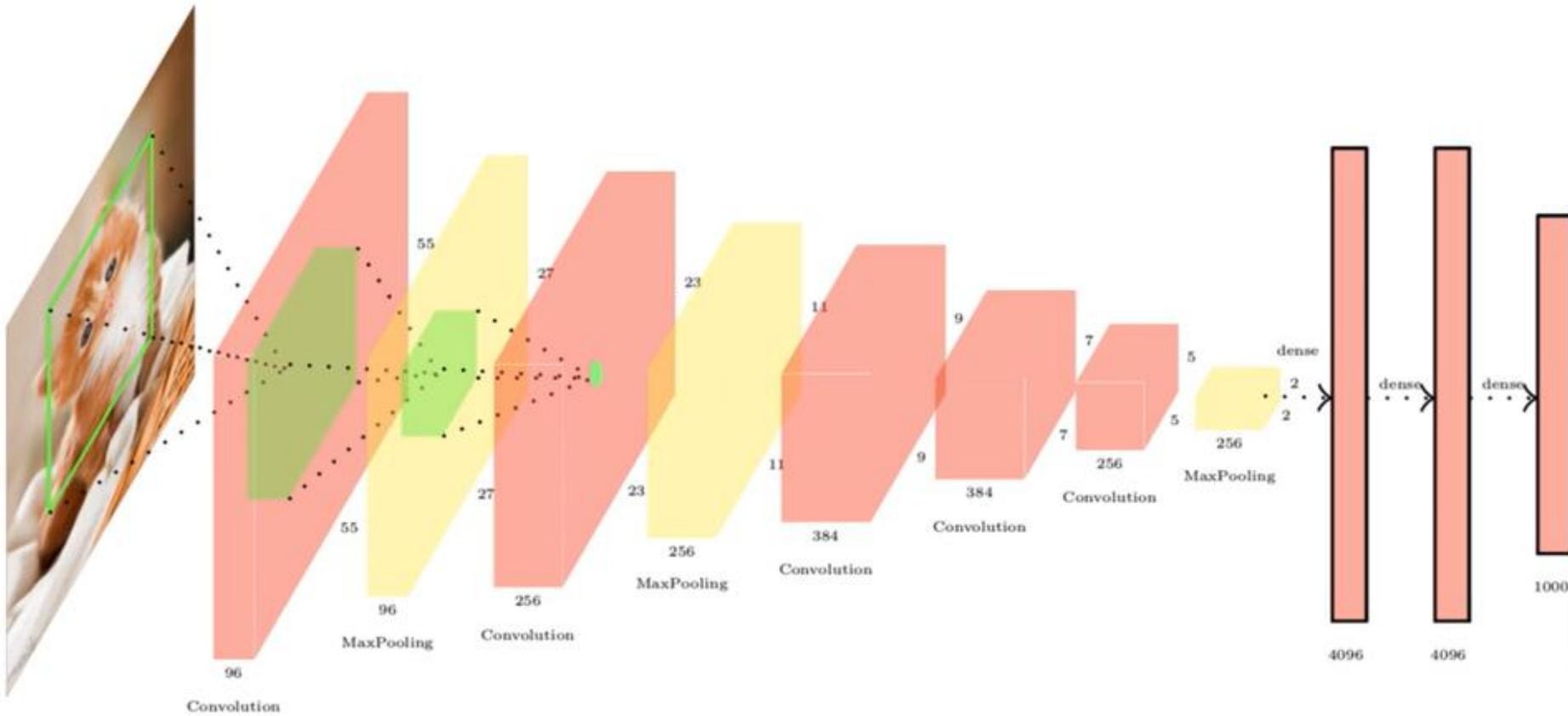
**ship**



**truck**

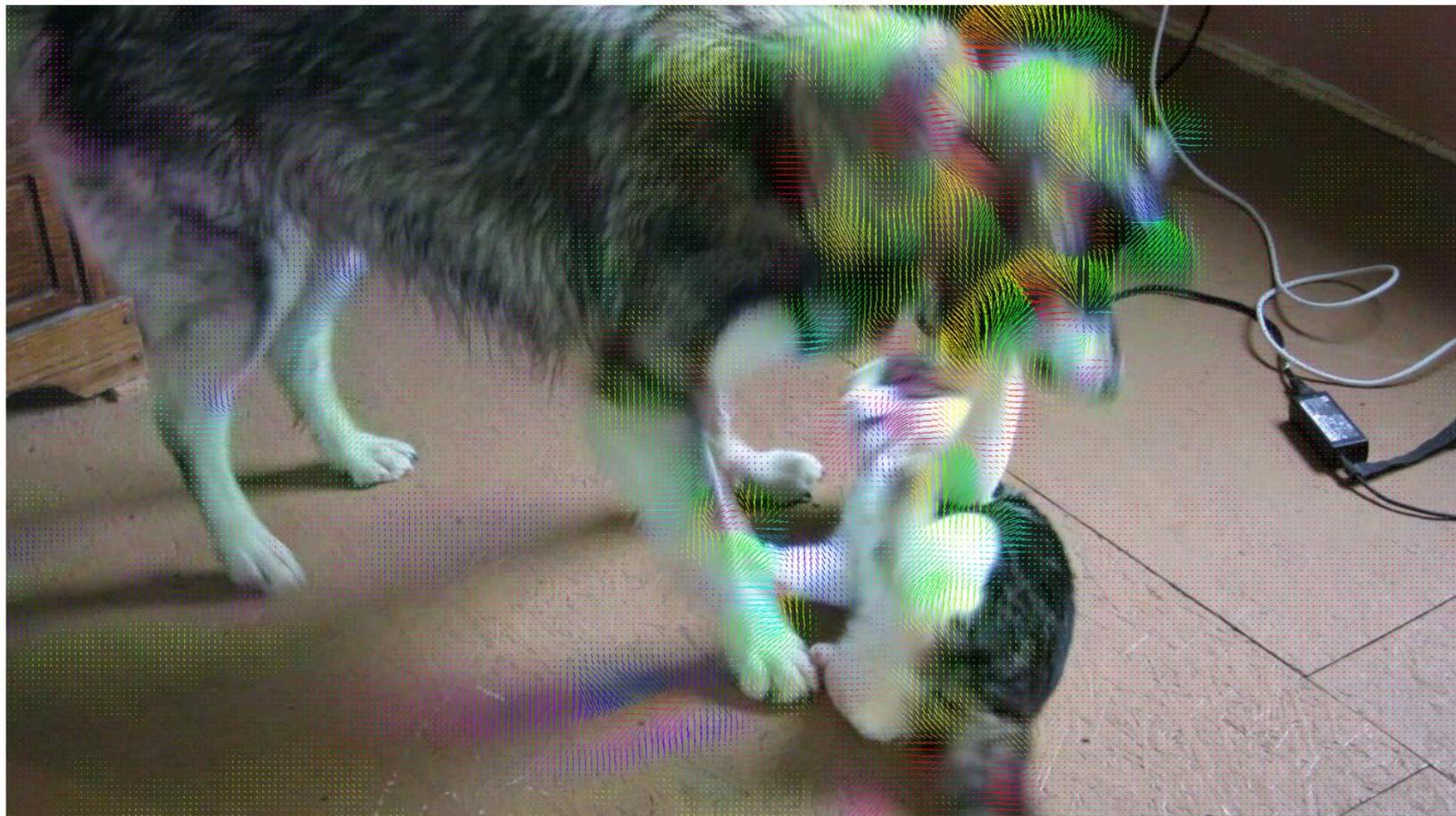


# Assignment 5: PyTorch

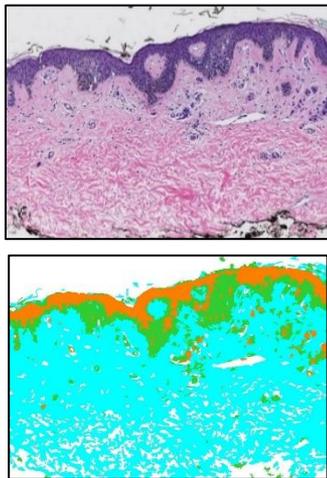
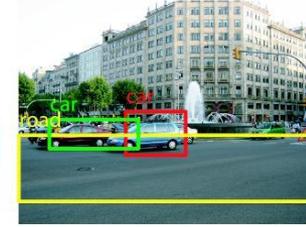
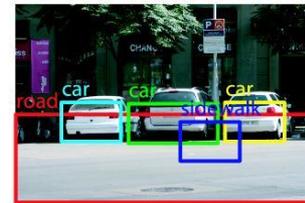
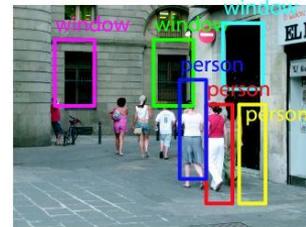
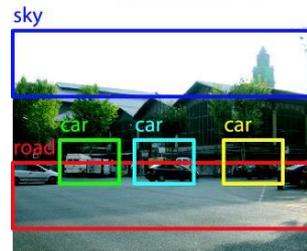
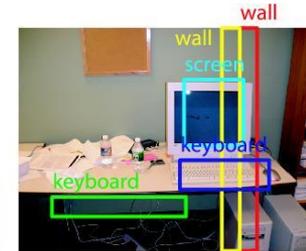
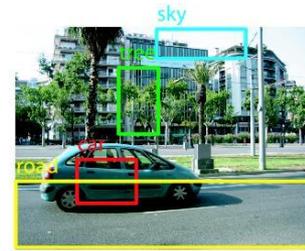
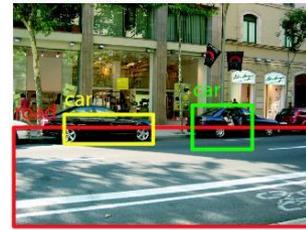
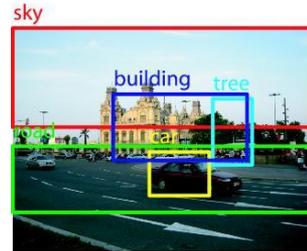


Input

# Assignment 6: Optical Flow

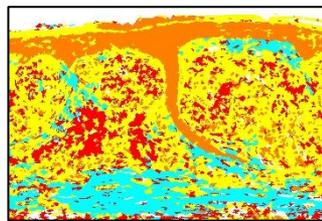
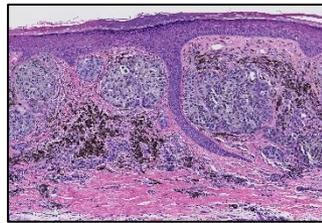


# Final Course Project: Machine Learning for Some Kind of Application

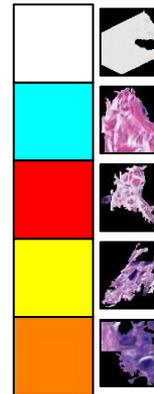


• b

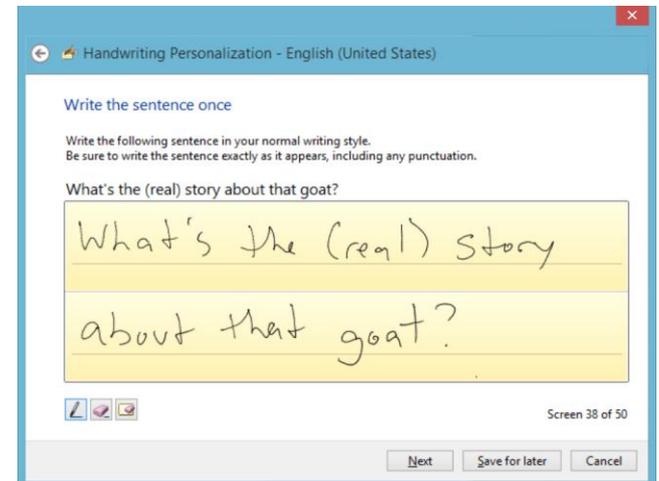
• a



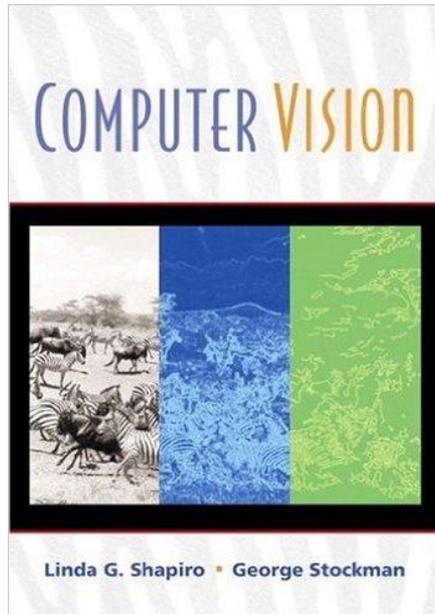
• c



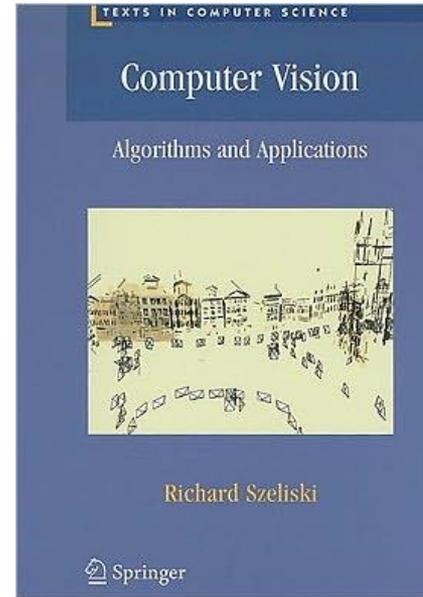
• d



# Books



Older, but designed for undergrads and has the basics. Chapters available from our web page.



Newest and available as a pdf online (2020 version).

# One Look Is Worth A Thousand Words--

One look at our line of Republic, Firestone, Miller and United States tires can tell you more than a hundred personal letters or advertisements.

WE WILL PROVE THEIR VALUE  
BEFORE YOU INVEST ONE DOLLAR  
IN THEM.

Ever consider buying Supplies from a catalog?

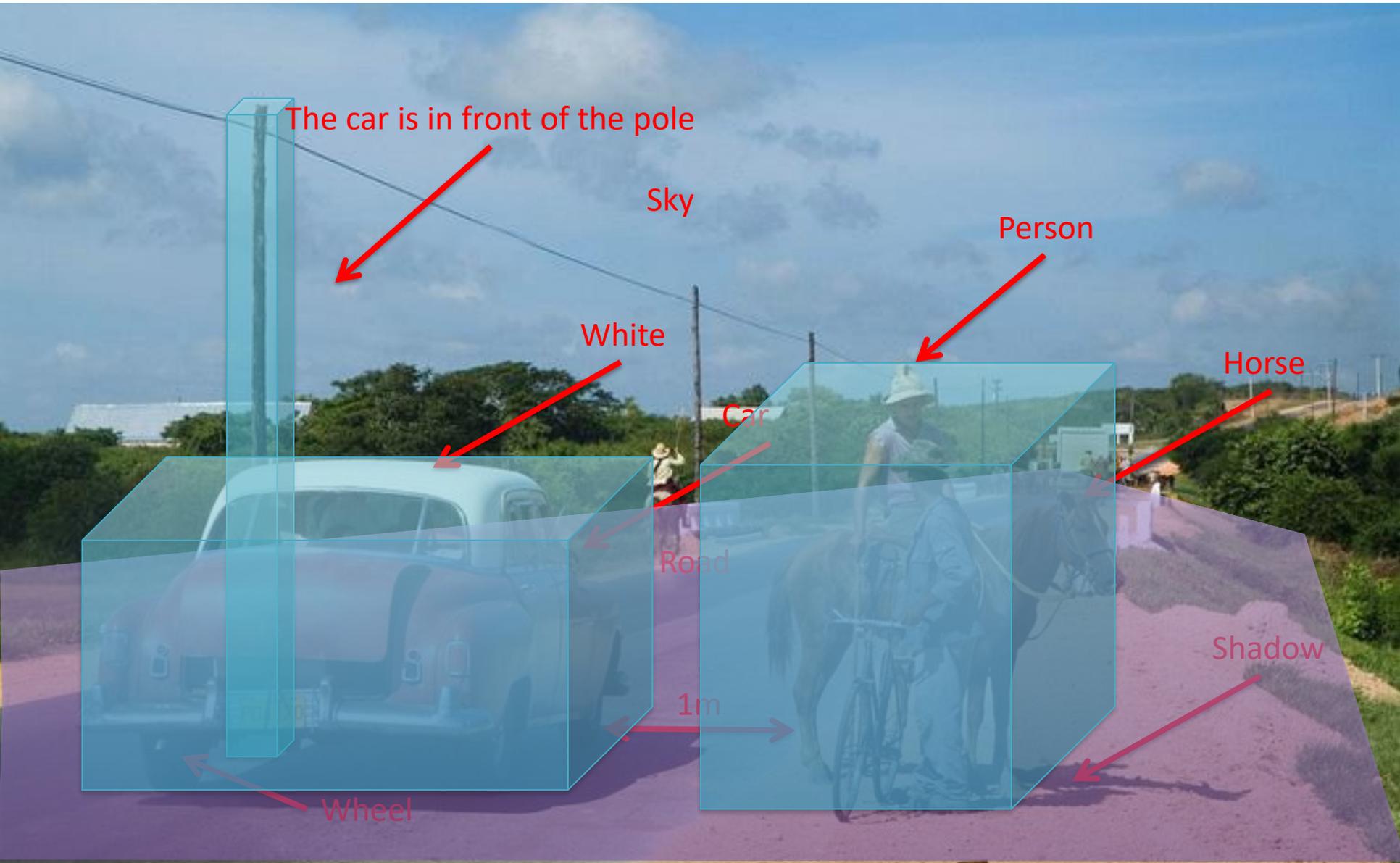
What's the use! Call and see what you are buying. One look at our display of automobile and motorcycle accessories will convince you of the fact.

THAT WE HAVE EVERYTHING FOR  
THE AUTO

## Piqua Auto Supply House

133 N. Main St.—Piqua, O.





The car is in front of the pole

Sky

Person

Horse

White

Car

Road

Shadow

Wheel

1m

# Computer Vision

- **Low Level Vision**
  - Measurements
  - Enhancements
  - Region segmentation
  - Features
- **Mid Level Vision**
  - Reconstruction
  - Depth
  - Motion Estimation
- **High Level Vision**
  - Category detection
  - Activity recognition
  - Deep understandings



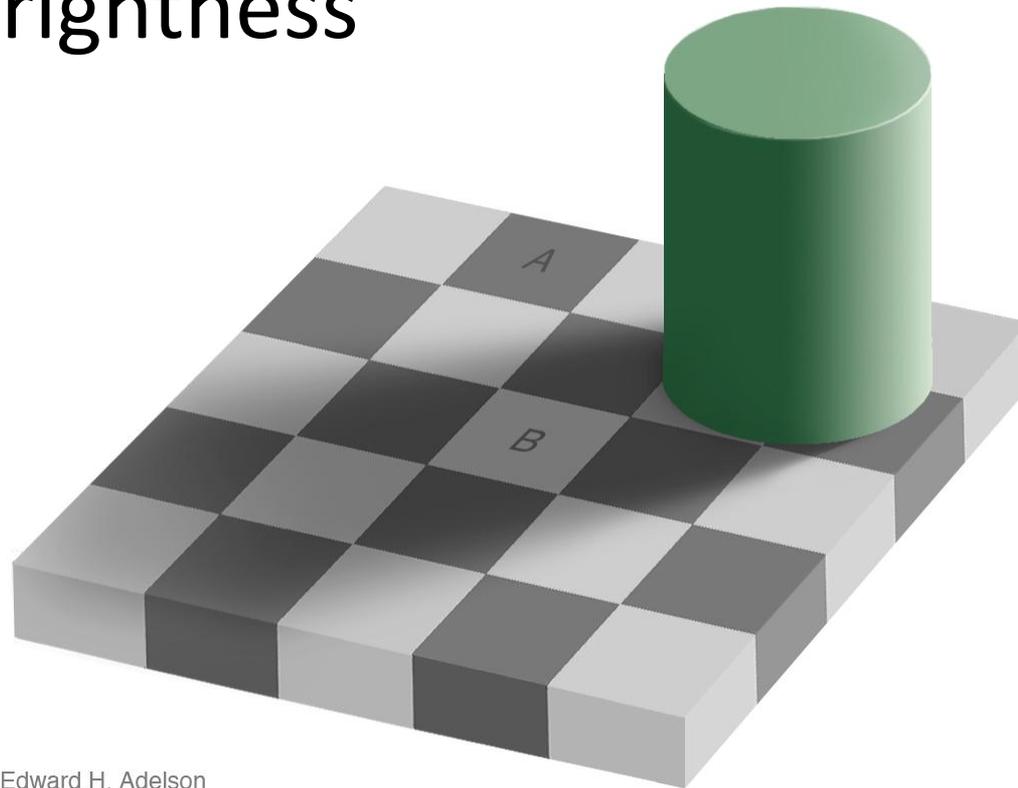
# Computer Vision

- Low Level Vision
  - Measurements
  - Enhancements
  - Region segmentation
  - Features
- Mid Level Vision
  - Reconstruction
  - Depth
  - Motion Estimation
- High Level Vision
  - Category detection
  - Activity recognition
  - Deep understandings



# Measurement

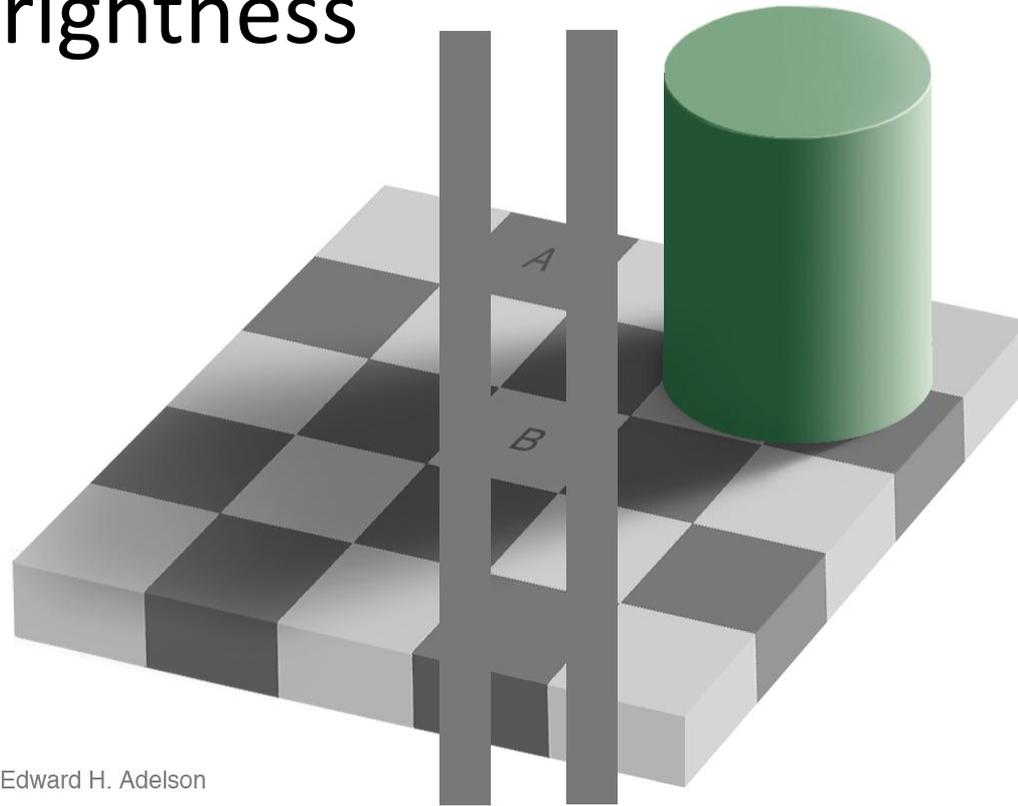
Brightness



Edward H. Adelson

# Measurement

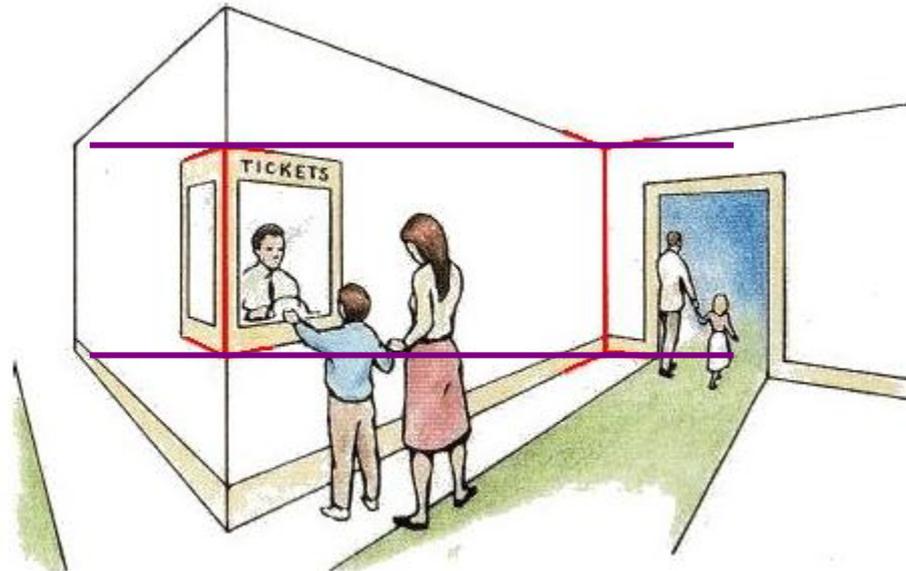
Brightness



Edward H. Adelson

# Measurement

## Length



Müller-Lyer Illusion

[http://www.michaelbach.de/ot/sze\\_muelue/index.html](http://www.michaelbach.de/ot/sze_muelue/index.html)

# Image Enhancement



*Image Inpainting*, M. Bertalmío et al.

<http://www.ia.upf.es/~mbertalmio//restoration.html>

# Image Enhancement



*Image Inpainting*, M. Bertalmío et al.

<http://www.ia.upf.es/~mbertalmio//restoration.html>

# Image Enhancement



*Image Inpainting*, M. Bertalmío et al.

<http://www.iaa.upf.es/~mbertalmio//restoration.html>

# Seam Carving

less  
important





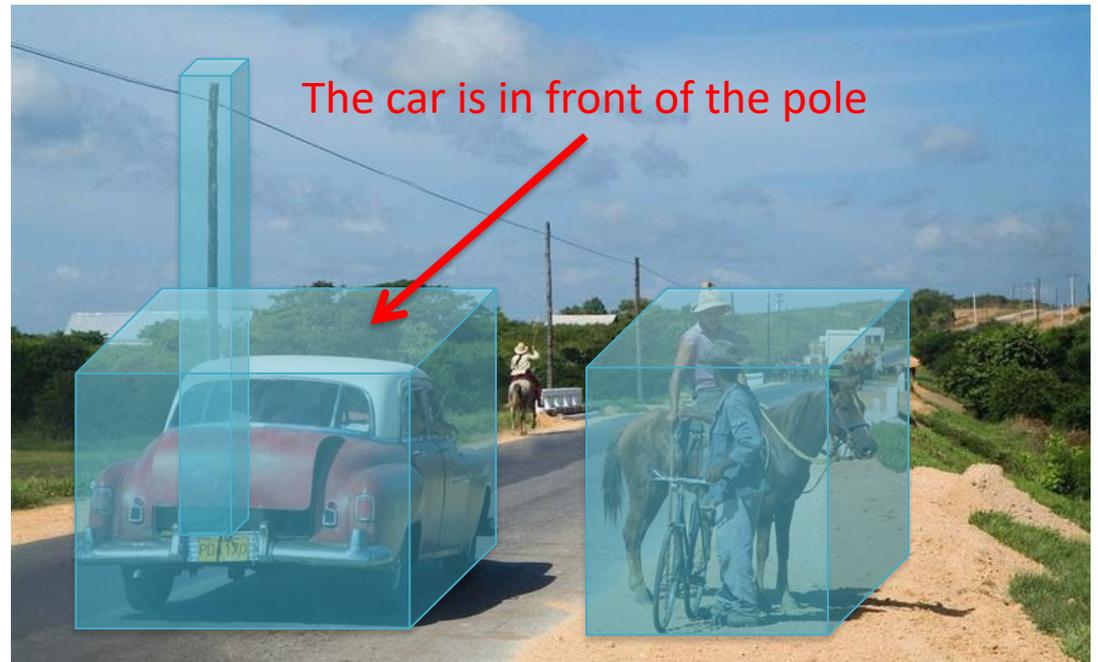
**Traditional resizing uses and stretches the whole image.**



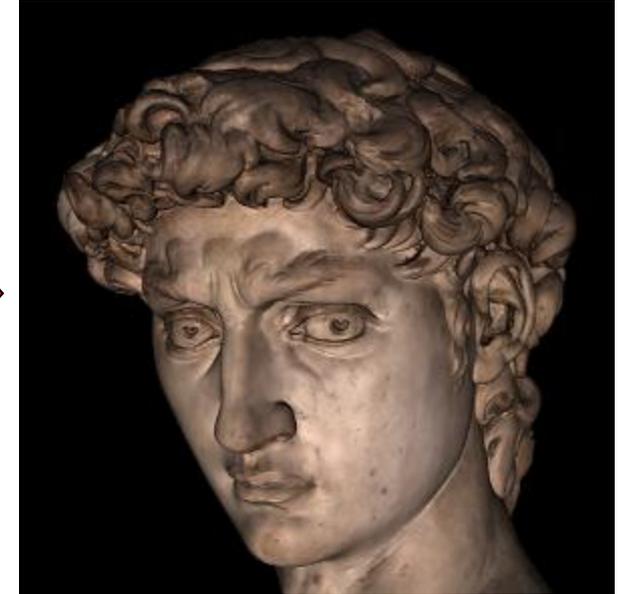
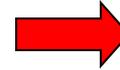
**Content-aware resizing uses important areas.  
Extends in horizontal direction and reduces in vertical.**

# Computer Vision

- Low Level Vision
  - Measurements
  - Enhancements
  - Region segmentation
  - Features
- Mid Level Vision
  - Reconstruction
  - Depth
  - Motion Estimation
- High Level Vision
  - Category detection
  - Activity recognition
  - Deep understandings



# Applications: 3D Scanning



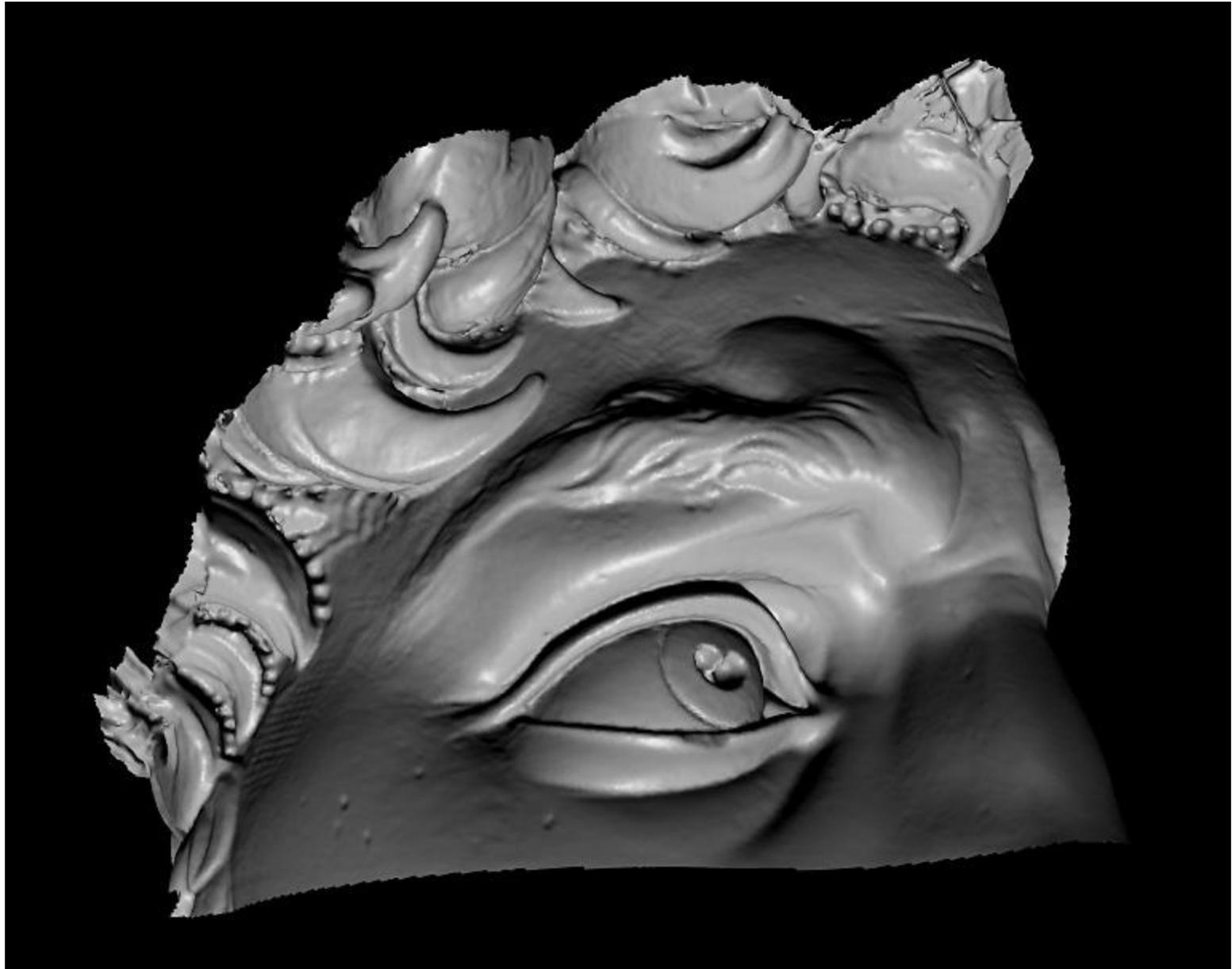
Scanning Michelangelo's "*The David*"

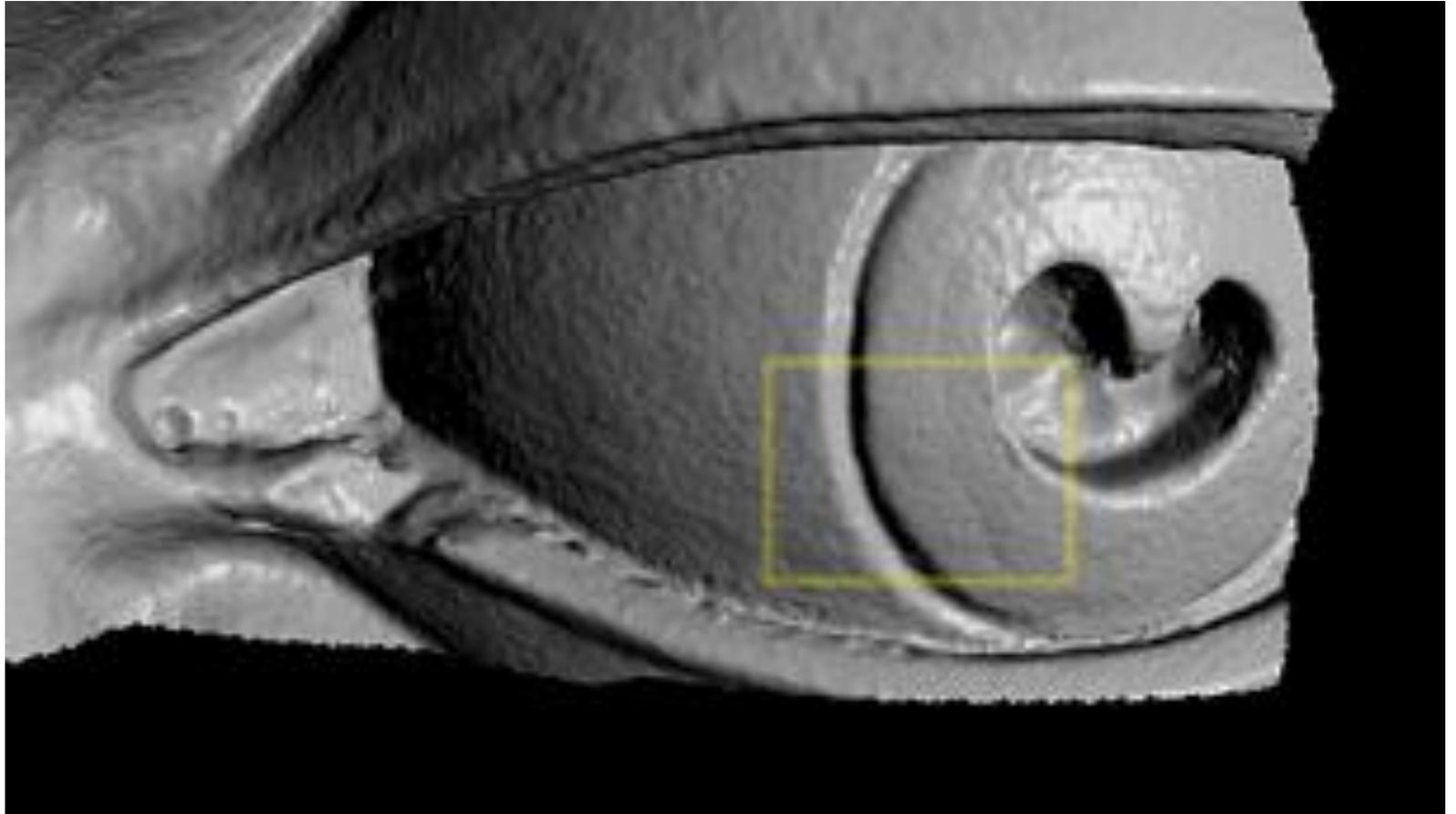
- [The Digital Michelangelo Project](http://graphics.stanford.edu/projects/mich/) (Marc Levoy)
  - <http://graphics.stanford.edu/projects/mich/>
- UW Prof. [Brian Curless](#), collaborator
- Linda's former student [Kari Pulli](#), collaborator
- 2 BILLION polygons, accuracy to .29mm

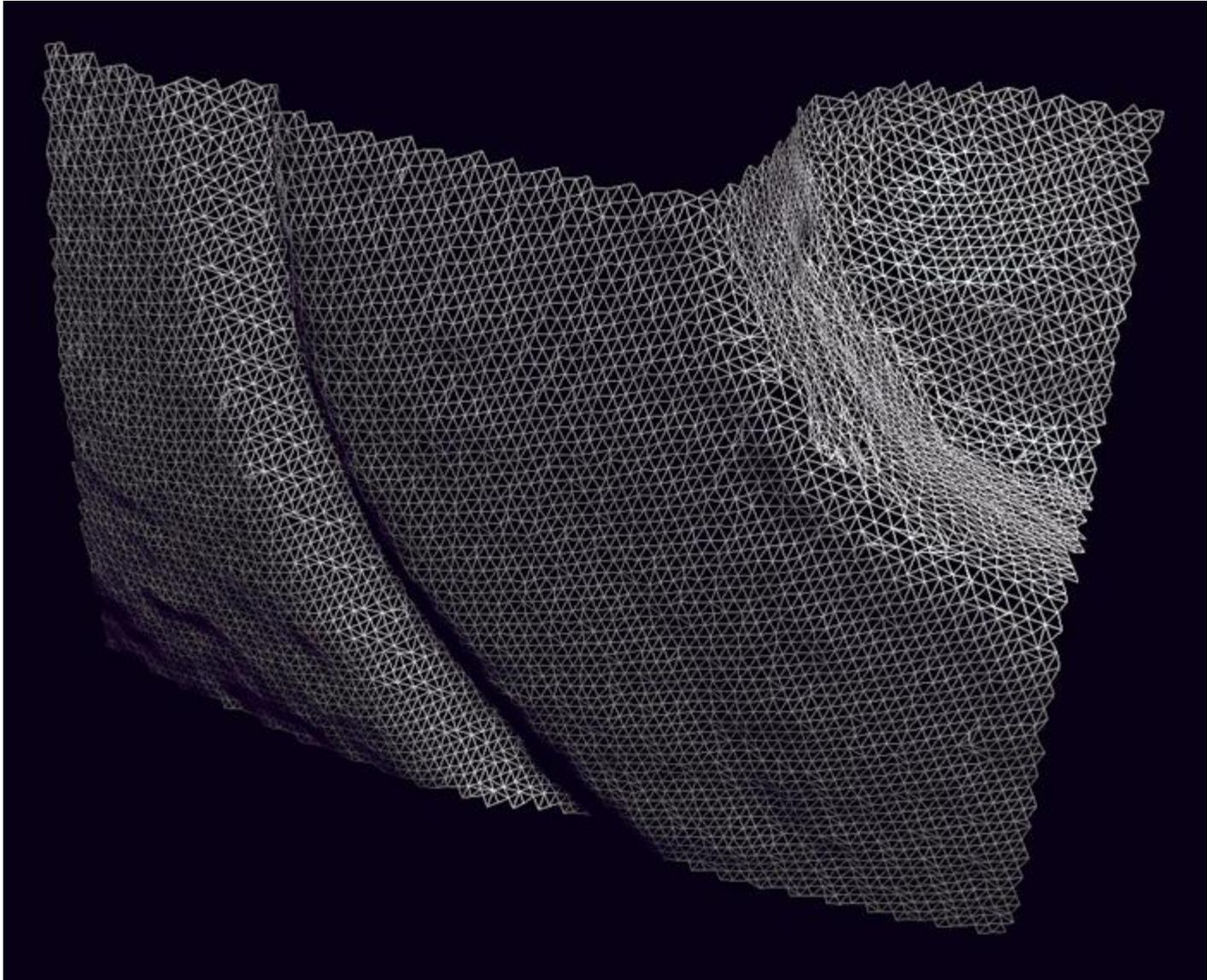


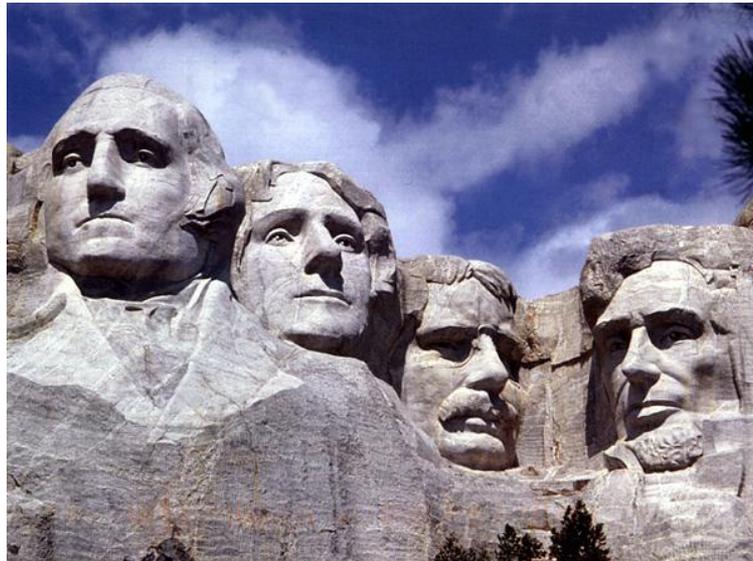
*The Digital Michelangelo Project, Levoy et al.*





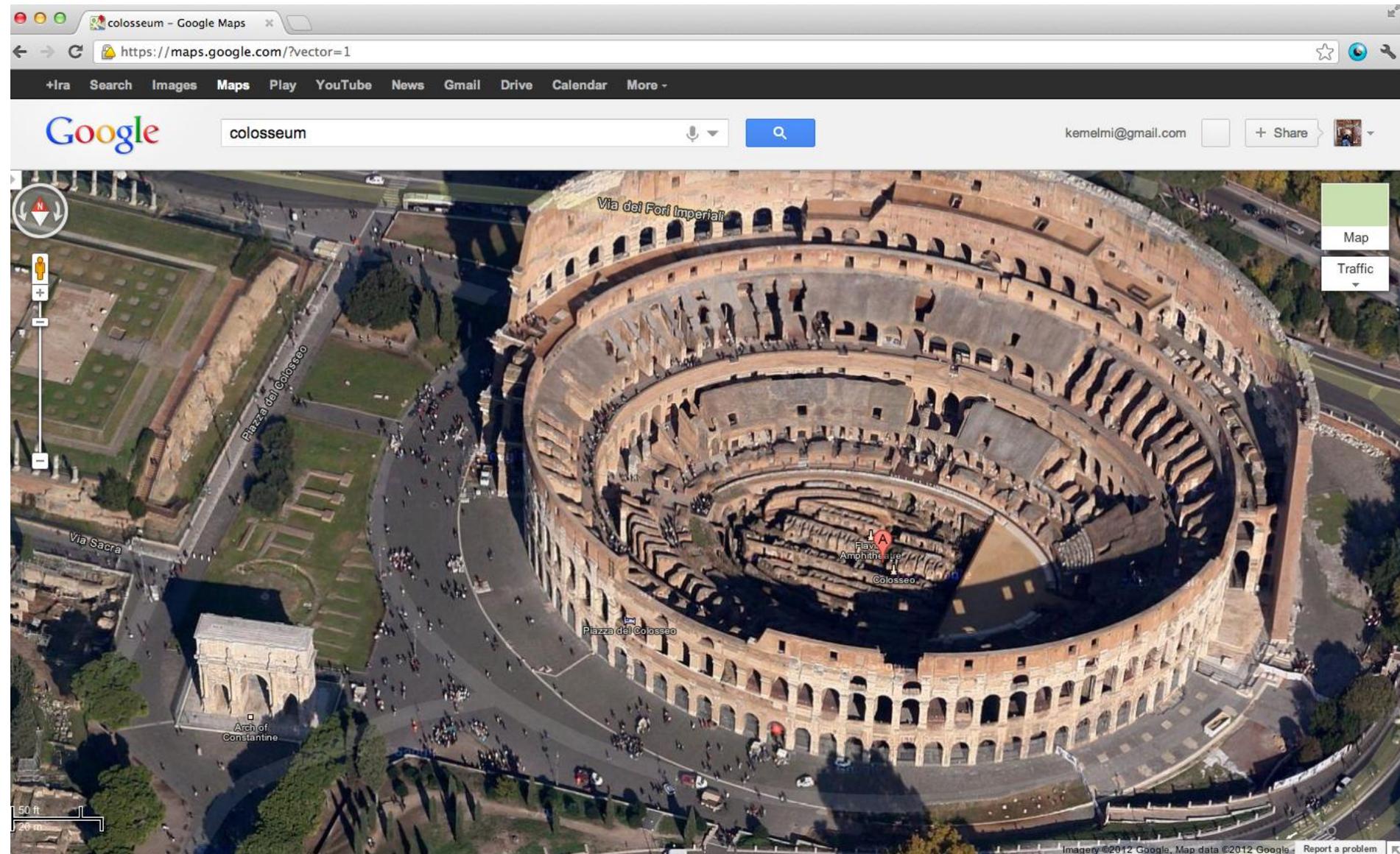






# Google's 3D Maps

## Structure estimation from tourist photos



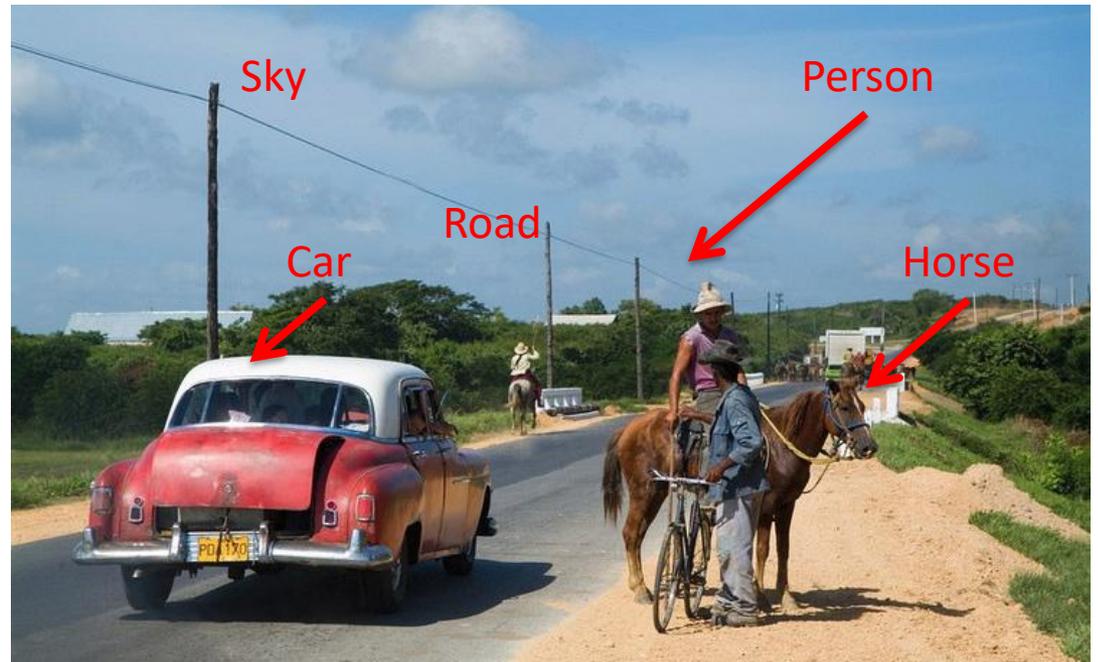
# Apple's 3D maps



<https://www.youtube.com/watch?v=InIVv-LsgZE>

# Computer Vision

- Low Level Vision
  - Measurements
  - Enhancements
  - Region segmentation
  - Features
- Mid Level Vision
  - Reconstruction
  - Depth
  - Motion Estimation
- High Level Vision
  - Category detection
  - Activity recognition
  - Deep understandings
  - Pose estimation

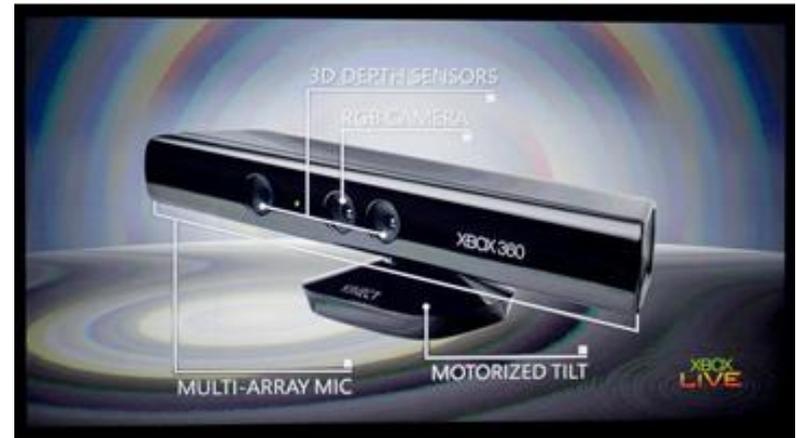


# Face detection

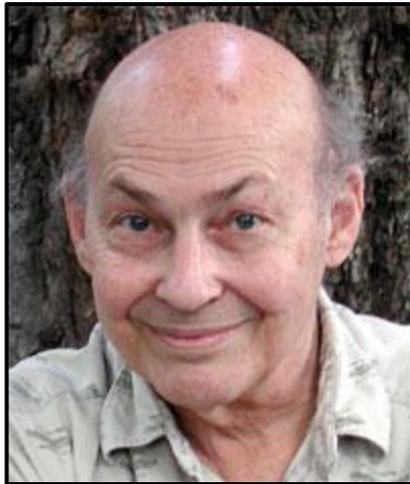


- Many new digital cameras now detect faces
  - Canon, Sony, Fuji, ...

# Vision-based interaction: Xbox Kinect



How hard is computer vision?



Marvin Minsky, MIT  
Turing award, 1969

“In 1966, Minsky hired a first-year undergraduate student and assigned him a problem to solve over the summer: connect a television camera to a computer and get the machine to describe what it sees.”

Crevier 1993, pg. 88

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

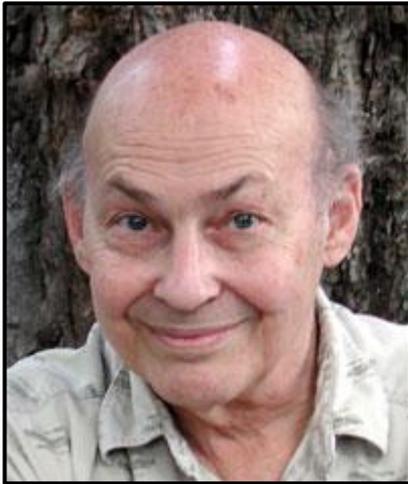
Artificial Intelligence Group  
Vision Memo. No. 100.

July 7, 1966

THE SUMMER VISION PROJECT

Seymour Papert

The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".



Marvin Minsky, MIT  
Turing award, 1969



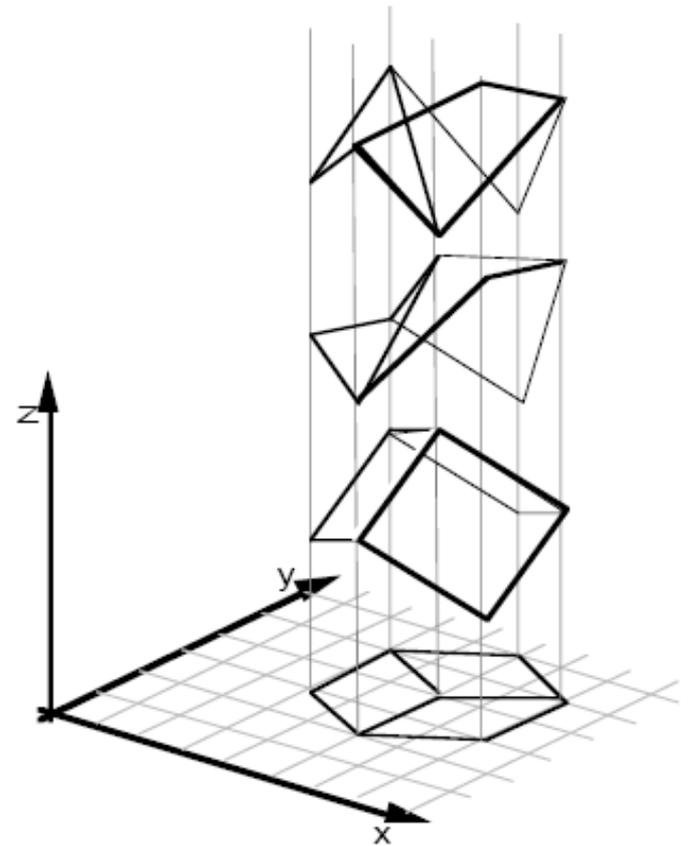
Gerald Sussman, MIT  
(the undergraduate)

“You’ll notice that Sussman never worked  
in vision again!” – Berthold Horn

Why vision is so hard?

# Why is vision so hard?

- Ill-posed problem



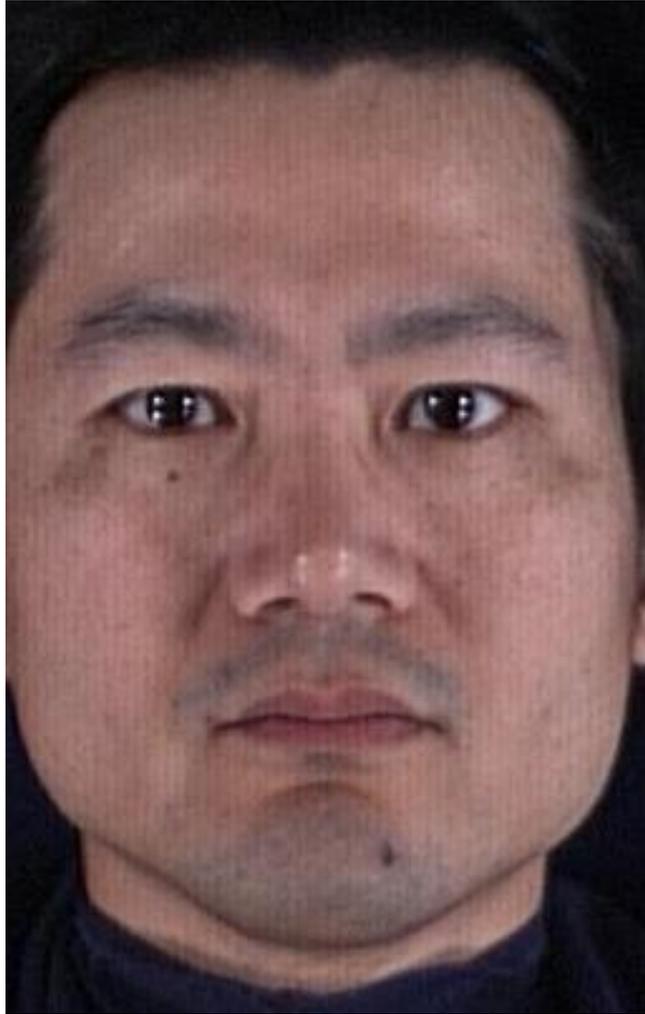
[Sinha and Adelson 1993]

# Challenges 1: view point variation



Michelangelo 1475-1564

## Challenges 2: illumination



# Challenges 3: occlusion

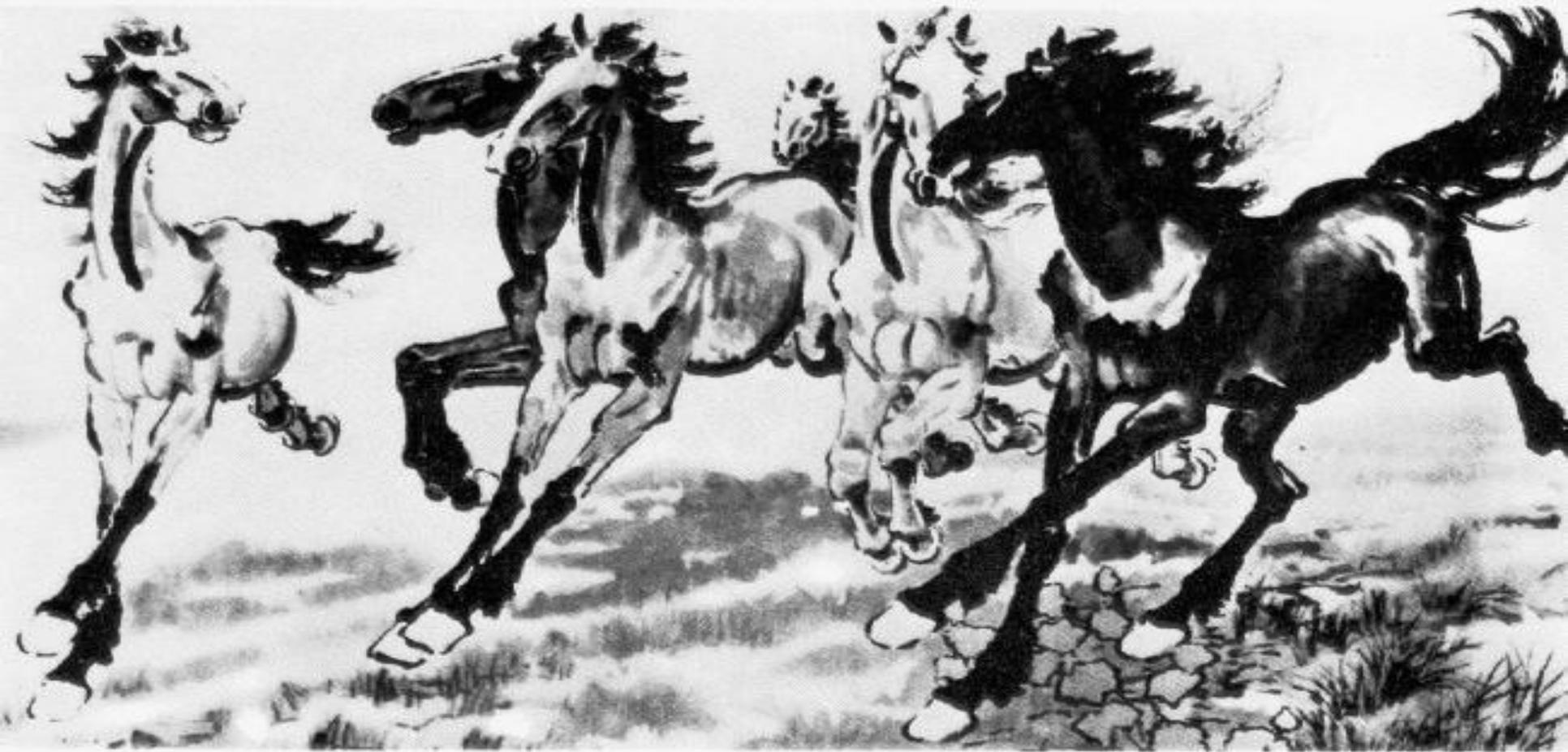


Magritte, 1957

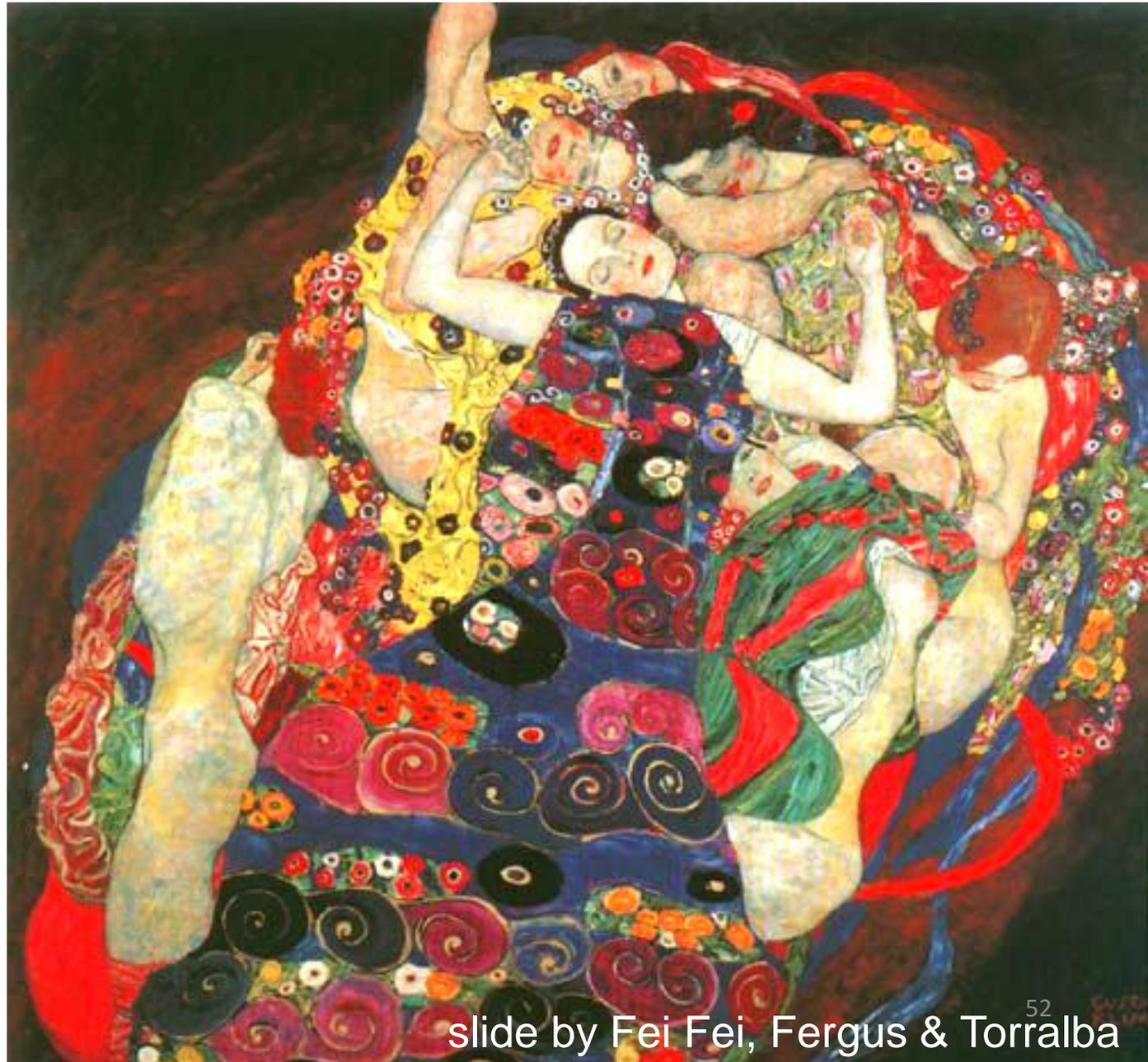
# Challenges 4: scale



# Challenges 5: deformation

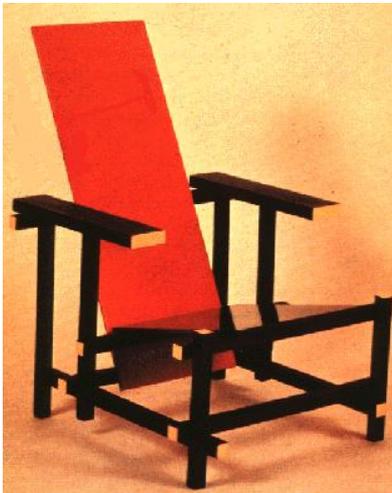


# Challenges 6: background clutter

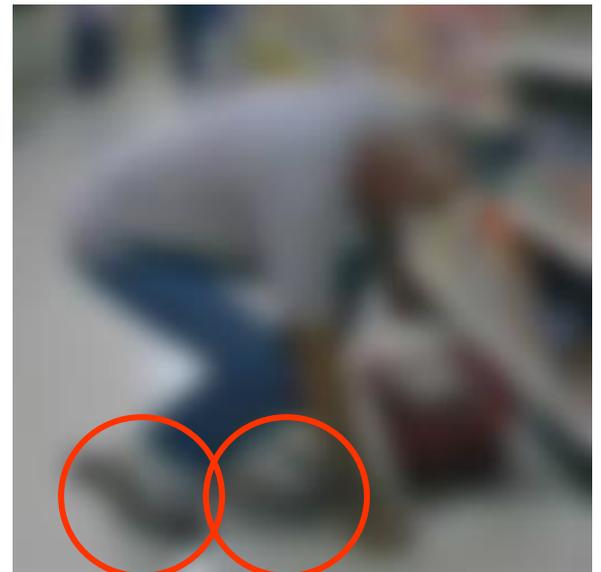
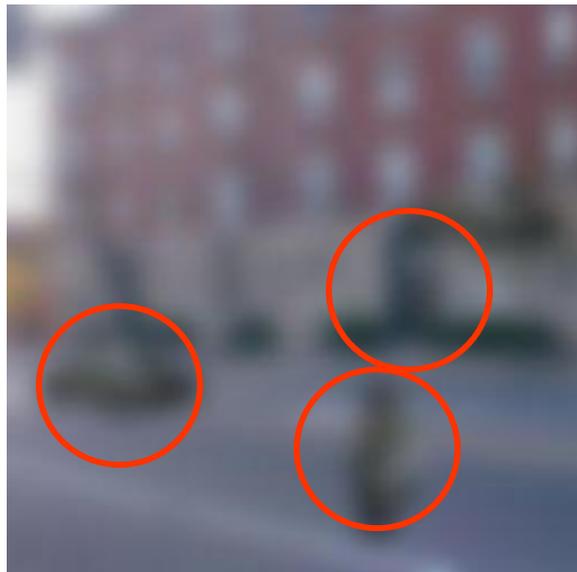
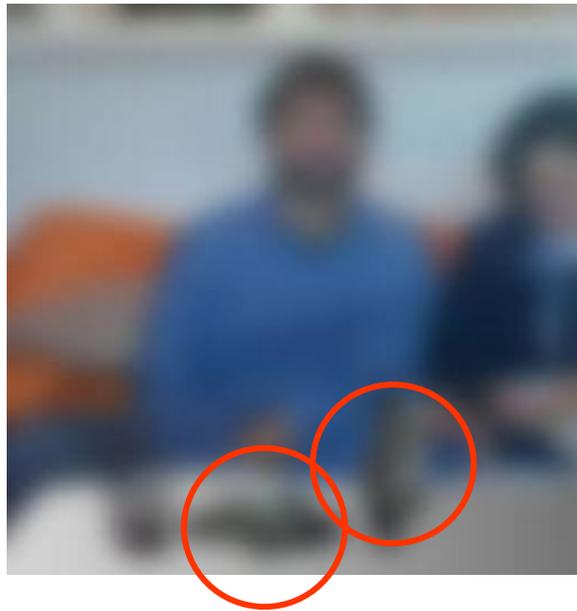


Klimt, 1913

# Challenges 7: object intra-class variation



# Challenges 8: local ambiguity



# Challenges 9: the world behind the image

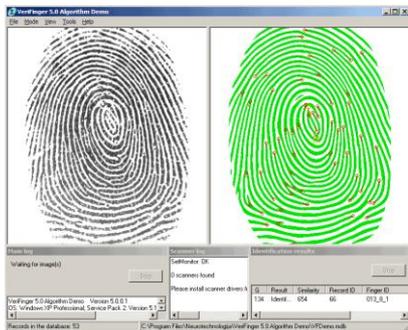


# What Works Today?

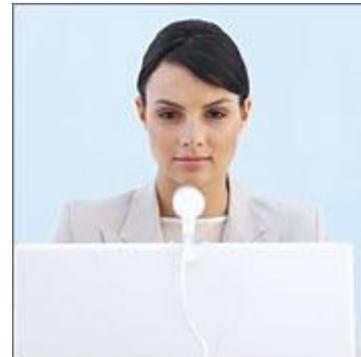
- Reading license plates, zip codes, checks

3 6 8 1 7 9 6 6 9 1  
6 7 5 7 8 6 3 4 8 5  
2 1 7 9 7 1 2 8 4 5  
4 8 1 9 0 1 8 8 9 4  
7 6 1 8 6 4 1 5 6 0  
7 5 9 2 6 5 8 1 9 7  
2 2 2 2 2 3 4 4 8 0  
0 2 3 8 0 7 3 8 5 7  
0 1 4 6 4 6 0 2 4 3  
7 1 2 8 7 6 9 8 6 1

# Biometrics



Fingerprint scanners on many new laptops, other devices



Face recognition systems now beginning to appear more widely

<http://www.sensiblevision.com/>

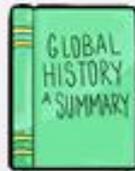
# Mobile visual search: Google Goggles

## Google Goggles in Action

Click the icons below to see the different ways Google Goggles can be used.



Landmark



Book



Contact Info.



Artwork



Places



Wine



Logo



# Face detection

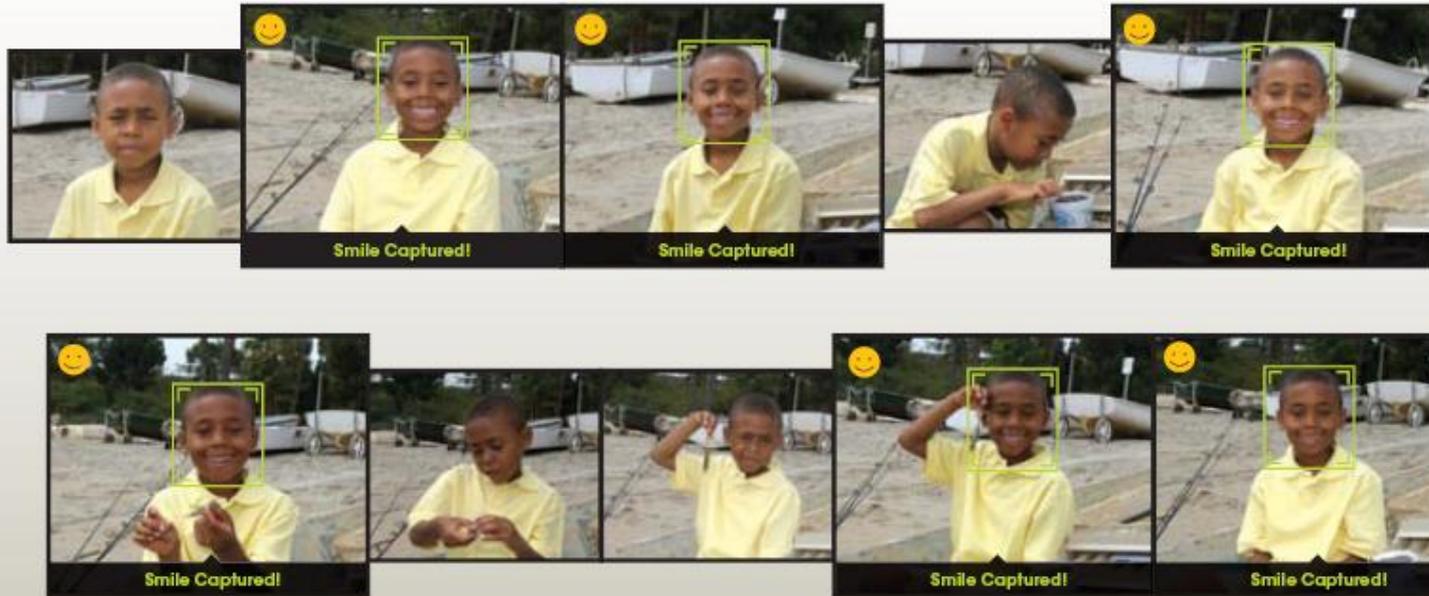


- Many digital cameras now detect faces
  - Canon, Sony, Fuji, ...

# Smile detection

## The Smile Shutter flow

Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Cyber-shot® camera can automatically trip the shutter at just the right instant to catch the perfect expression.



# Face recognition: Apple iPhoto, Facebook, Google, etc



# Object recognition (in supermarkets)



## [LaneHawk by EvolutionRobotics](#)

“A smart camera is flush-mounted in the checkout lane, continuously watching for items. When an item is detected and recognized, the cashier verifies the quantity of items that were found under the basket, and continues to close the transaction. The item can remain under the basket, and with LaneHawk, you are assured to get paid for it... “



**News Front Page**



**Africa**

**Americas**

**Asia-Pacific**

**Europe**

**Middle East**

**South Asia**

**UK**

England

Northern Ireland

Scotland

**Wales**

UK Politics

Education

Magazine

**Business**

**Health**

**Science &**

Last Updated: Wednesday, 31 August 2005, 05:44 GMT 06:44 UK

[✉ E-mail this to a friend](#)

[🖨️ Printable version](#)

## Computer alert for drowning girl

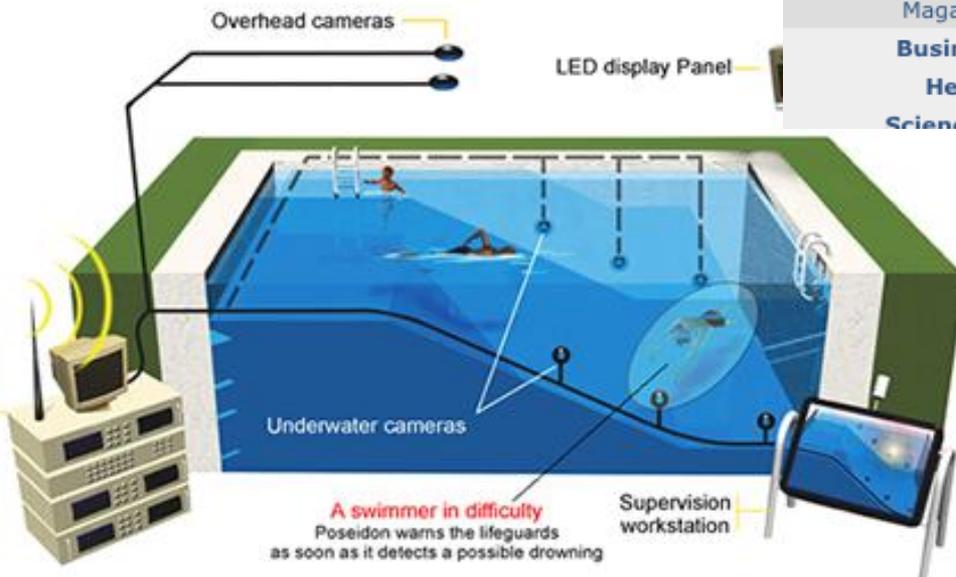
**A 10-year-old girl has been saved from drowning by a computer system designed to raise the alarm when swimmers get into difficulties.**



[▶ VIDEO](#) **Watch the rescue**

The girl, from Rochdale, was at the deep end of the pool in Bangor, north Wales, when she sank to the bottom.

The £65,000 system, called Poseidon, detected her on the pool floor and sounded the alarm. A lifeguard pulled her out and she recovered in hospital.



# Security

Local 

## Cameras help confirm Scott suicide ruling

Friday, December 04, 2009



TAGS: [local](#), [paul meincke](#)

 [Comment Now](#) [Email](#) [Print](#) [Report a typo](#) 



**Paul Meincke**

More: [Bio](#), [News Team](#)

December 4, 2009 (CHICAGO) (WLS) -- Chicago police have closed the case in the death of Chicago School Board President Michael Scott.

Police Supt. Jody Weis says investigators used police cameras in the city to trace Scott's last steps in the hours before his body was found in November.

Scott's death has been ruled a suicide. The medical examiner's office concluded --not long after Scott's body was found -- that he had committed suicide. Police did not dispute the finding but wanted to pursue all the investigative leads they could. They say they have done that and have now reached the same conclusion.

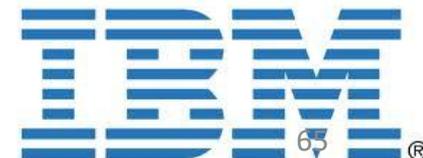
### Share this Story

 [Recommend](#)  Be the first of your friends to recommend this.

 [Tweet](#)   [+1](#) Recommend this on Google

[News Headlines](#)  [Video](#) 

-  2 suspects arrested in volleyball star's murder **47 min ago**
-  BP Gas Recall: BP finds, fixes source of bad gas
-  Teachers union, board resume negotiating
-  Back to School
-  5 injured in South Side shooting **49 min ago**
-  Pastor: Stacy Peterson said she lied for Drew



# Automotive safety

manufacturer products    consumer products

## Our Vision. Your Safety.

rear looking camera    forward looking camera

side looking camera

› **EyeQ** Vision on a Chip

› **Vision Applications**  
Road, Vehicle, Pedestrian Protection and more

› **AWS** Advance Warning System

› **News**

› Mobileye Advanced Technologies Power Volvo Cars World First Collision Warning With Auto Brake System

› Volvo: New Collision Warning with Auto Brake Helps Prevent Rear-end

› all news

› **Events**

› Mobileye at Equip Auto, Paris, France

› Mobileye at SEMA, Las Vegas, NV

› read more

- [Mobileye](#): Vision systems in high-end BMW, GM, Volvo models
  - Pedestrian collision warning
  - Forward collision warning
  - Lane departure warning
  - Headway monitoring and warning

# Google cars



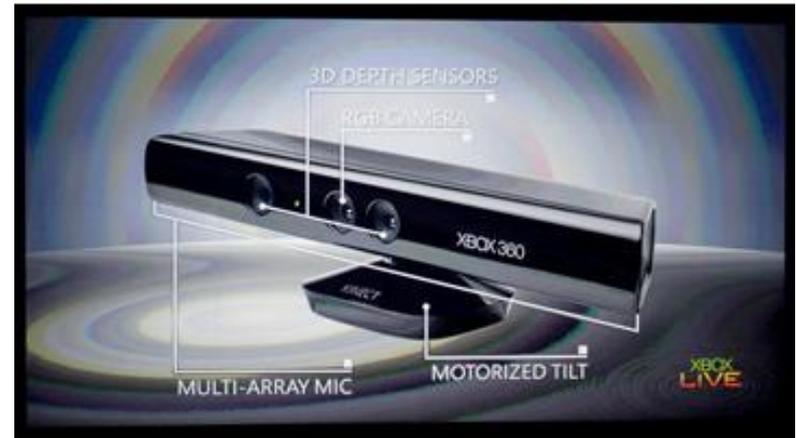
Oct 9, 2010. ["Google Cars Drive Themselves, in Traffic"](#). *The New York Times*. John Markoff

June 24, 2011. ["Nevada state law paves the way for driverless cars"](#). *Financial Post*.

Christine Dobby

Aug 9, 2011, ["Human error blamed after Google's driverless car sparks five-vehicle crash"](#). *The Star* (Toronto)

# Vision-based interaction: Xbox Kinect



# Augmented reality, consumer products



Mark Billinghamurst and Tom Furness developed with me.

# Special effects: shape and motion capture



# Vision for robotics, space exploration

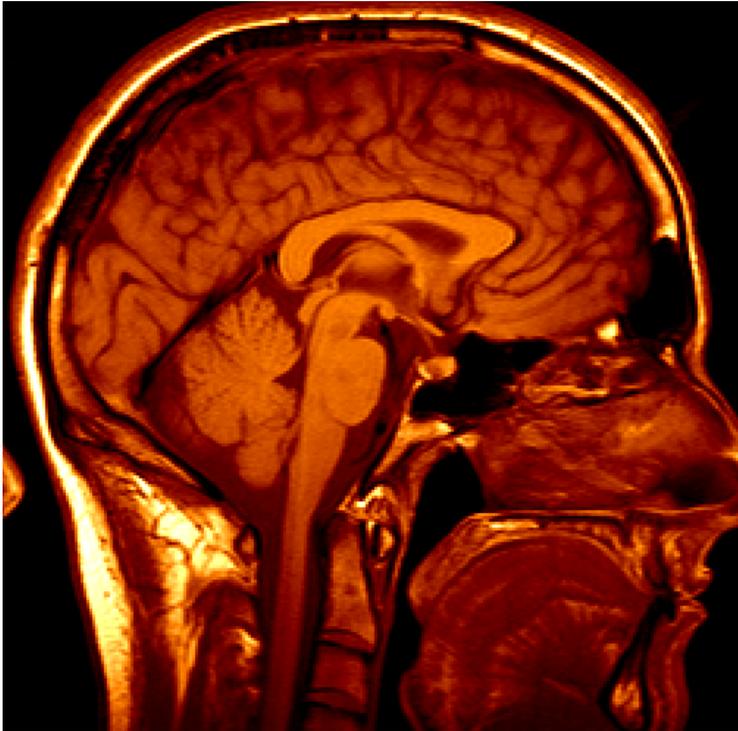


[NASA'S Mars Exploration Rover Spirit](#) captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

## Vision systems (JPL) used for several tasks

- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking

# Medical imaging

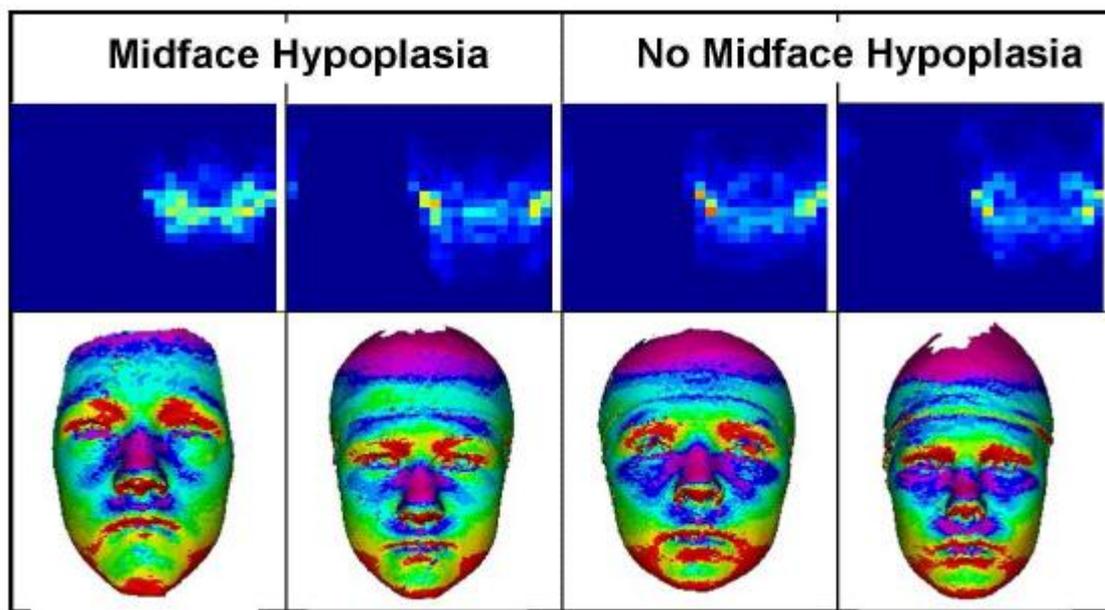


3D imaging  
MRI, CT



Image guided surgery  
[Grimson et al., MIT](#)

# Classification of 22q11.2DS



- Treat 2D azimuth-elevation angle histogram as feature vector

|               | 8×8   | 16×16 | 24×24 | 32 × 32 | Experts' median |
|---------------|-------|-------|-------|---------|-----------------|
| Whole 2D hist | 0.651 | 0.569 | 0.79  | 0.684   | 0.68            |

# Computer vision research in healthcare



assisted living, patient monitoring  
[Lan et al, PAMI 2012]



autism screening

<http://www.gatech.edu/newsroom/release.html?nid=60509>

# Computer vision in the real-world

- Most examples are less than 7 years old
- Very active research area. Many new applications to come.
- A website of computer vision industries maintained by Prof. David Lowe (UBC)

<http://www.cs.ubc.ca/~lowe/vision.html>

- Note: website is old but interesting
- Note: David Lowe retired and moved to Google 2015 to 2018

# Assignments

- Assignment 1: Fun with Color
- Assignment 2: Image Resizing and Filtering
- Assignment 3: Panorama Stitching
- Assignment 4: Neural Networks
- Assignment 5: Pytorch
- Assignment 6: Optical Flow
- Course Project: Teams working on Machine Learning Projects

# Assignment 1

- It's about color, which we will cover Wednesday.
- It's meant to be very easy, but you want to start it early.

# Assignment 1 Parts

- 1. data class for an image

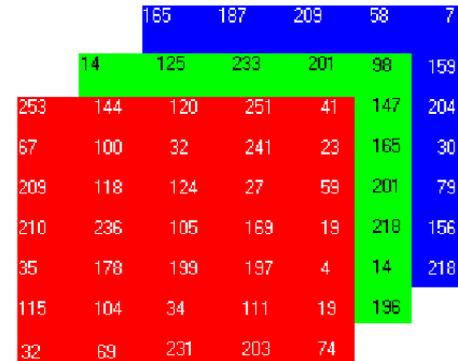
```
class Image:
```

```
    w: int
```

```
    h: int
```

```
    c: int
```

```
    data: np.ndarray
```



- So an image is a 3D array with channels, height, and width (like for colors).
- data is floating point numbers between 0 and 1

# Assignment 1: Parts

## Read them

- TODO #1: `get_pixel` and `set_pixel`
- TODO #2: `copy_image`
- TODO #3: `rgb_to_grayscale`
- TODO #4: `shift_image` (shifts values)
- TODO #5: `clamp_image` (get values between 0 and 1)
- TODO #6: `rgb_to_hsv`
- TODO #7: `hsv_to_rgb`

Have Fun